SEEDS

Working Paper Five:

Data Service Provider Model,

Requirements and Levels of Service

April 24, 2002

G. Hunolt, SGT, Inc.

Outline

1 Introduction

This working paper is the fifth of a set of papers that describes the SEEDS (Strategic Evolution of Earth Science Enterprise Data Systems) Levels of Service (LOS) / Cost Estimation (LOS/CE) study. The study goal is to develop a cost estimation model and coupled requirements and levels of services to support the SEEDS Formulation team in estimating the life cycle costs of future ESE data service providers and supporting systems, where 'data service provider' is used as a generic term for any data/information related activity. The set of working papers is intended to serve as a vehicle for coordinating work on the study, obtaining feedback and guidance from ESDIS SOO and the user community, and as embryos of reports that will be produced as the task proceeds.

As working papers, each version of each paper that appears represents a snapshot in time, with the work in various stages of completion; as work progresses the content (and sometimes the organization) of the working papers will change reflecting progress made, responses to feedback and guidance received, etc.

This fifth working paper of the set describes the requirements and levels of services component of the general data service provider reference model developed for the LOS/CE study, and will reflects results of the February, 2002, SEEDS Community Workshop.

1.1 Levels of Service

A major objective of the LOS/CE study is to assist the SEEDS Formulation Team in establishing the minimum (and recommended) levels of service for ESE data service providers. These levels of service will be refined in a bottoms-up manner through community workshops of potential providers and users of ESE data services. To facilitate this process, a user-oriented view of the levels of service is included in this paper.

Levels of service will be associated with functional requirements, describing different degrees of performance with which the requirement would be met. For example, a functional requirement might be: "The data service provider shall distribute data and products to users on media". Accompanying this requirement might be descriptions of quantitatively distinct levels of service, such as "delivery on media shall be provided within one working day of receipt of a data request", "delivery on media shall be provided within two calendar weeks of receipt of a data request", and "delivery on media shall be provided within one calendar month of receipt of a data request". Which level of service would be most appropriate ('recommended') or acceptable ('minimum') for a particular ESE data service provider would depend on its particular mission and the needs of its users. Not all requirements have levels of service associated with them; by their nature, some requirements are either met or not met without any shades of gray.

The levels of service and their associated requirements will also feed into the life cycle cost estimation phase of the study because data service provider costs must be driven by / associated with the levels of service required of the data service provider. Successful development of a life cycle cost estimation capability will be dependent on an accurate assessment of the levels of services needed from ESE data service providers.

The requirements developed by this study are not intended to serve as the complete definition of the requirements side of a contract between an ESE program office and ESE data service providers, or to serve as a basis for procurements. This study ignores questions about what level of requirements will be 'owned' at the program level vs by data service providers themselves.

1.2 Levels of Service and the Data Service Provider Reference Model

The requirements / levels of service are one of three related aspects of the Data Service Provider Reference Model, a general functional model of a generic data service provider:

1) A set of 'functional areas' that collectively comprise the full range of functions that a data service provider might perform and the areas of cost that must be considered by the cost estimation by analogy model. The functional areas of the reference model are defined in Working Paper 3, "Data Service Provider Reference Model - Functional Areas".

2) A set of parameters for each functional area that constitute a quantitative description of the workload, staff effort, and any other factors that contribute to cost for that area, additional 'roll-up' parameters that sum items such as staff effort across the functional areas, and other parameters like labor rates that are required for cost estimation. The parameters of the reference model are defined in detail in Working Paper 4, "Data Service Provider Reference Model - Model Parameters".

3) A set of requirements and levels of service for each functional area. These are defined in Section 2 of this working paper.

These three aspects of the model are closely coupled to ensure the internal consistency of the model. The set of functional areas is the underpinning; both the model parameters and requirements / levels of service are organized according to the functional areas. The requirements / levels of service and the model parameters are coupled in that the definitions of the requirements / levels of service embody model parameters. This integration of the three aspects of the model is intended to ensure that estimated costs are driven by and traceable to requirements to the fullest extent possible. Working Paper 4 includes a mapping of the Data Service Provider Reference Model parameters to the requirements / levels of service. The intent is to show which parameters fall within the scope of each requirement, and to ensure that each requirement / levels of service that should have one or more parameters associated with it actually does.

The intent of the requirements / levels of service described in Section 2 below (and the corresponding functional areas presented in Working Paper 3) is to provide a complete description at a reasonable level of detail of the abstract ESE data service provider, and to reflect the concerns expressed in the February, 2002, SEEDS community workshop. The ability of the cost estimation by analogy approach to reflect the full range of detail described in the functional areas and requirements / levels of service will be limited by the information available in the comparables database and the feasibility of reasonable assumptions where information is not available. This will be reflected in the reference model's parameter set, described in Working Paper 4.

As the needs of the ESE science and applications program evolve, and hence the ESE roles and missions for data service providers evolve, and as information technology that touches all aspects of every data service provider and the user community evolves (e.g. the GRID distributed computing approach), the data service provider requirements and levels of service will evolve. The content of this paper can only represent a snapshot in time - and indeed a snapshot that is in part tied to current and recent past experience with working data service providers. If the cost estimation tool (and the underling data service provider model) proves useful, it will have to be maintained and revised perhaps dramatically to preserve or improve its usefulness over time.

In addition to evolving with changing ESE program needs, the cost estimation by analogy model (and the data service provider model) will be improved in successive iterations as the comparables database grows and includes more new activities, and with lessons learned derived from use of earlier versions of the model.

Section 2 presents the requirements / levels of service organized by the reference model's functional areas. Section 3 presents a user-oriented view of the levels of service (with requirements implied rather than stated explicitly).

Appendix A includes a draft set of program level requirements, "NewDISS Level 0 Requirements", GSFC, September 2001, which were used as a starting point / umbrella for the requirements / levels of service in Section 3.

2 Data Service Provider Requirements / Levels of Service

This section presents the requirements / levels of service template for a generic ESE data service provider, organized by the data service provider reference model's fourteen functional areas. As such it does not imply or embody any architecture, i.e. any allocation of requirements to various particular components.

The term 'template' is used for two reasons. The first is that all of the requirements / levels of service will not apply to all actual ESE data service providers. The second is that the requirements contain placeholders for specifics that must be filled in (i.e. choices between alternatives shown, or between possible levels of service, or replacement of placeholders with lists or numerical values) to generate from the template a set of requirements / levels of service that would apply to a specific ESE data service provider, and that would allow a cost estimate for it to be produced.

Appendix A contains a draft set of high level or programmatic requirements referred to as "NewDISS Level 0 Requirements" produced by the SEEDS Formulation Team in September, 2001. These provide an "umbrella" for the requirements described below. Additional guidance for the initial set of requirements and levels of service was drawn from the ESDIS Project Level 2 Requirements for EOSDIS Version 0, updated March 2000, which addressed requirements and levels of service, the report "Global Change Science Requirements for Long-Term Archiving", NOAA-NASA and USGCRP Program Office, March 1999, and the report "Ensuring the Climate Record from the NPP and NPOESS Meteorological Satellites", NRC Committee on Earth Studies, September 2000.

The requirements and levels of service were updated following the February, 2002, SEEDS community workshop, responding to comments and recommendations made at the workshop and in white papers submitted to the workshop.

Placeholders for items to be specified when the template is to be used to generate requirements for a specific data service provider are enclosed in brackets […].

2.1 Ingest Requirements / Levels of Service

a. The data service provider shall ingest the following data [ingest data stream table, listing for each data stream: name, source, product types ingested, product type format (input and conversion after ingest if any) products ingested per day of each type, volume ingested per day]. The input data streams should cover all data to be received by the center, e.g. satellite data streams, ancillary data products, processed products generated by other data service providers, etc., based on its ESE mission, and accompanying metadata, documentation, retention plan (e.g. a part of a life cycle data management plan) etc.

   Levels of Service:

   1) operational (time-critical) ingest with immediate verification of data integrity and quality;

   2) routine ingest and verification of data quality and integrity without tight time constraints;

   3) ad hoc or intermittent ingest on a non-operational basis with verification of data quality and integrity;

   4) ad hoc or intermittent ingest on a non-operational basis.

   Levels of service can be mixed within a data service provider; i.e. different levels may be appropriate for different data streams.

b. The data service provider shall provide standard metrics on ingest to the SEEDS Office.

2.2 Processing Requirements / Levels of Service

a. The data service provider shall generate the following products ('standard products' characterized by a peer reviewed, validated, reasonably stable, 'science quality' processing algorithm), included required Level 1B products [standard product table, listing for each product type/series: name, format, retention

plan, product instances produced per day, volume per day, required input data streams] on a highly reliable, operational basis, either on a routine schedule or on-demand, based on its ESE mission.

Levels of Service:

1) operational products shall be generated within 2 days of ingest/availability of required inputs;

2) operational products shall be generated within 7 days of ingest/availability of required inputs;

3) operational products shall be generated within 30 days of ingest/availability of required inputs.

b. The data service provider shall generate the following products [product table, listing for each product type/series: name, format, retention plan, average product instances produced per day, average volume per day, required input data streams] on an ad hoc, non-operational basis. (The product table can refer to known or expected products, or can be used to establish a capacity to support a level of ad hoc product generation (perhaps data mining or data integration) that will be used to support user needs as they arise.)

Levels of Service:

1) specific targets for processing adopted on a case by case basis;

2) general goals for processing;

3) no goals, purely ad hoc processing.

c. The data service provider shall provide a capacity for reprocessing of standard products [standard product table] on an ad hoc basis in response to reprocessing requests.

1) the aggregate capacity for reprocessing shall be 9 times the original aggregate processing rate;

2) the aggregate capacity for reprocessing shall be 6 times the original aggregate processing rate;

3) the aggregate capacity for reprocessing shall be 3 times the original aggregate processing rate.

d. The data service provider shall reprocess standard products [standard product table, listing for each product a reprocessing interval] according to a reprocessing schedule.

Levels of Service:

1) reprocessing shall be performed according to a negotiated reprocessing schedule;

2) reprocessing shall be performed to meet the general goals of a nominal schedule;

3) reprocessing shall be performed following a nominal schedule on a resource / time available basis.

e. The data service provider shall accept science algorithm software from users for [product list], and perform integration and test of the software, and operational execution of the software to produce products.

1) the data service provider shall accept standard, research product generation software, and/or data integration and data mining software from users;

2) the data service provider shall accept research product generation software and/or data integration and data mining software from users;

3) the data service provider shall accept standard and/or research product generation software from users;

4) the data service provider shall accept research product generation software from users;

5) the data service provider shall accept standard product generation software from users.

f. The data services provider shall be capable of cross-calibration of data from multiple sources to produce consistent product time series spanning multiple instruments / platforms.

g. The data service provider shall provide standard metrics on production to the SEEDS Office.

2.3 Documentation Requirements / Levels of Service

a. The data service provider shall generate and provide ESE/ SEEDS adopted standards compliant catalog information (metadata, including browse) and documentation describing all data and information produced and/or acquired and held by the data service provider.

   Levels of Service:

   1) data and product holdings (including multiple versions of products and corresponding documentation as needed) documented to the ESE / SEEDS adopted standard for long term archiving, including details of processing algorithms, processing history, many etc.;

   2) documentation ensured to be sufficient for current use (e.g. product type descriptions, product instance (a.k.a. granule) descriptions including version information, FAQs, 'readme's, web pages with links to metadata, user guides, references to journal articles describing the production or use of the data or product);

   3) documentation only as received from product provider.

b. The data service provider shall update documentation of data and products with user comments, e.g. on parameter accuracy, product usability, data services available or needed for a product, etc.

   Levels of Service:

   1) data and products routinely updated with user comments;

   2) data and products occasionally updated with user comments;

   3) data and products rarely updated with user comments.

c. The data service provider shall generate and provide DIF (Directory Interchange Format) documents to the Global Change Master Directory on all products available from the data service provider prior to their release for distribution.

2.4 Archive Requirements / Levels of Service

a. The data service provider shall add to its archive or working storage the following data and products [archive product table, drawn from ingest data stream table, standard product, and ad hoc product tables and reprocessing volume] and related documentation / metadata.

b. The data service provider shall provide for secure, permanent storage of data at the "raw" sensor level (NASA Level 0 plus appended calibration and geolocation information).

c. The data service provider shall provide for secure storage of all standard or other science products it produces until the end of the science mission or until transfer to an approved permanent archive, per the applicable life-cycle data management plan (or separate retention plans).

d. The data service provider shall have the capability to selectively replace archived product instances (single or large sets) with new versions, and to selectively update metadata and documentation (e.g. to update quality flags when a product is validated).

e. The data service provider shall provide for an [archive] [working storage] capacity of [number] TB.

   Levels of Service:

   1) archive capacity is cumulative sum of all data ingested plus all products generated (including allowance for retaining multiple versions of the same product as required to provide needed support to the provider's science or applications community);

   2) archive capacity is limited to a specified threshold.

f. The data service provider shall perform quality screening on data entering the archive (e.g. read after write check when data is written to archive media) and exiting the archive (e.g. tracking of read failures and corrected errors or other indication of media degradation on all reads from archive media).
Levels of service:

1) exit and entry screening;

2) entry screening.

g. The data service provider shall take steps to ensure the preservation of data in its archive.
Levels of service:

1) 10% per year random screening to detect and replace failing / degrading media;

2) 5% per year random screening;

3) 1% per year random screening.

h. The data service provider shall provide a backup and restore capability for its [archive] [working storage].
Levels of service:

1) full off-site backup, with regular sampling and exercise of restore capability to verify integrity;

2) partial, [Backup Fraction - % of archive backed up], off-site backup, with sampling;

3) partial, [Backup Fraction - % of archive backed up], on-site backup, with sampling.

i. The data service provider shall use robust archive media.
Levels of Service:

1) archive media consistent with best commercial practice;

2) archive media and system vendor independent;

3) archive media vendor independent.

j. The data service provider shall plan and perform periodic migration of archive to new archive media / technology.
Levels of Service:

1) planned and budgeted for migration;

2) no planned migration, but ad hoc migration as need is seen to arise.

(Note - this requirement would not apply to a data service provider with a shorter lifetime than a migration cycle appropriate for its archive media / technology.)

k. The data service provider shall provide standard metrics on archive to the SEEDS Office.

2.5 Search and Order Requirements / Levels of Service

a. The data service provider shall provide users with access to all metadata and information holdings.
Levels of Service:

1) public access to all users;

2) access to the science and applications community;

3) access to a limited team of scientists or applications specialists.

b. The data service provider shall provide a world wide web accessible search and order capability to [all users (including the general public) consistent with SEEDS standards and practices] [a limited set of science team members]. (Scope consistent with the level of service for requirement 2.5 a above.)
Levels of Service:

1) allow search for instances of multiple product types that pertain to a specified object or phenomenon (e.g. a named hurricane, a volcanic eruption, a field campaign, etc.)

2) allow search for instances of multiple product types by geophysical parameter(s), time, and space applied across multiple product types;

3) allow search for instances of multiple product types by common time and space criteria (coincident search);

4) allow search for instances of single product type by time and space criteria;

5) allow search for particular instances of a product type from a list of those available.

c. The data service provider shall provide the user with the option of quickly viewing information describing any product returned as meeting search criteria.

Levels of Service:

1) descriptive information includes detailed algorithm and use explanations, references to a few published papers that describe the production or use of the product, standard guide and DIF metadata.

2) descriptive information includes references to a few published papers that describe the production or use of the product, standard guide and DIF metadata.

3) descriptive information includes standard guide and DIF metadata.

d. The data service provider shall provide an interface for system-system search and order access as well as an interface for human users.

e. The data service provider shall provide an interface to and support selected external catalog search capabilities (e.g. EDG, Mercury, Echo).

2.6 Access and Distribution Requirements / Levels of Service

a. The data service provider shall provide users with access to all data and product holdings, including all standard science products (Level 1b, Level 2, and Level 3) produced by the data service provider.

Levels of Service:

1) public access to all users;

2) access to the science community or an applications community;

3) access to a limited team of scientists or applications specialists.

b. The data service provider shall provide data and products to users in (at a minimum) one of the SEEDS core formats.

c. The data service provider shall enhance its distribution capability with supporting data services such as subsetting, resampling, reformatting (e.g. to GIS formats), reprojecting and/or packaging to meet the needs of its users.

Levels of Service:

1) supporting data services available for most archived data and products;

2) supporting data services available for less than half of archived data and products;

3) supporting data services available for a few selected data and products only.

The particular supporting services available would vary on a product by product basis, depending on the nature of the product and the needs of the user community.

d. The data service provider shall provide data to users on an [operational, subscription (i.e. standing order), and/or in response to request] basis. (An operational basis means in part that a data service provider will formally commit in a level of service agreement or equivalent to terms of service.)

e. The data service provider shall provide an interface for system to system network delivery of data and products.

f. The data service provider shall perform timely distribution of data and products to users by network, providing an average distribution volume capacity of [number] TB per day.

Levels of service:

1) availability of a single product for access by user software within ten seconds;

2) availability of a single product for network delivery (e.g. FTP pickup or push) within ten seconds;

3) availability of a single product for network delivery within ten minutes;

4) availability of a single product for network delivery within twenty four hours.

g. The data service provider shall perform timely distribution of data and products to users on SEEDS standard media types in response to user requests, providing an average volume capacity of [number] TB per day.

Levels of Service:

1) shipping of media product within three days of receipt of request;

2) shipping of media product within one week of receipt of request,

3) shipping of media product within one month of receipt of request.

h. The data service provider shall have the capacity to distribute products on an average of [number] media units per day.

i. The data service provider with final ESE archive responsibility (i.e., a Backbone Data Center unless, for example, or a Science Data Service Provider which has held its products to the time for their transfer to the long term archive) shall transfer its data, products, and documentation (done to the long term archive standard) to the designated long term archive according to its Life Cycle Data Management Plan.

j. The data service provider shall provide SEEDS standard metrics on distribution to the SEEDS Office.

2.7 User Support Requirements / Levels of Service

a. The data service provider shall be capable of supporting [number] of distinct, active users per year who request and/or access and use data service provider products.

1) one user support staff member per 100 active users;

2) one user support staff member per 500 active users;

3) one user support staff member per 1,000 active users.

(The number of active users is the number of distinct users who request, or through an automated means obtain, delivery of data and/or information products per year.)

b. The data service provider shall provide a trained user support staff.

Levels of service:

1) below plus science expertise in data / product quality and their research uses.

2) below plus technical expertise in data structures, use of tools for format conversions, subsetting, analysis, etc.

3) below plus comprehensive knowledge of details of formats for most if not all products;

4) user support staff are knowledgeable about the data service provider's holdings and ordering/delivery options.

(Not all members of the user support staff would necessarily have the highest level of expertise.)

c. The data service provider shall provide a help desk function (i.e., staff awaiting user contacts who can assist in ordering, track and status pending requests, resolve problems, etc.).
   Levels of Service:
   1) Help desk staffed seven days per week, twenty-four hours per day.
   2) Help desk staffed five days per week, twelve hours per day;
   3) Help desk staffed five days per week, eight hours per day;

d. The data service provider shall provide on-line user support (FAQ, data, product and service descriptions, etc.).

e. The data service provider shall perform user outreach, education, and training.
   Levels of Service:
   1) Below plus provide user training sessions at universities, schools, etc.
   2) Below plus expanded booth support including mini-workshops, user training sessions;
   3) Below plus booth support at four conferences per year;
   4) Produce and make available outreach material - pamphlets, brochures, posters, etc.

2.8 Instrument / Mission Operations Requirements / Levels of Service

a. The data service provider shall monitor the status and performance of [name] instruments and in some cases also [name] spacecraft for which it is responsible, generating instrument commands and in some cases spacecraft commands as needed.

b. The data service provider shall obtain the services of a NASA (or other spacecraft operator as appropriate) mission operations facility to provide instrument and spacecraft data and to receive, validate, and transmit instrument and/or spacecraft commands to the spacecraft.

2.9 Sustaining Engineering Requirements / Levels of Service

a. The data service provider shall maintain and, as needed, enhance custom software it develops to meet its mission needs, and reused software it customizes and integrates, a total of [number] SLOC.
   Levels of Service:
   1) no or very infrequent interruptions of data service provider operations;
   2) occasional interruptions in data service provider operations;
   3) as needed, with interruptions in data service provider operations a secondary concern.

2.10 Engineering Support Requirements / Levels of Service

a. The data service provider shall perform system administration, network administration, database administration, coordination of hardware maintenance by vendors, and other technical functions as required for performance of its mission.
   Levels of Service:
   1) no or very infrequent interruptions of data service provider operations;
   2) occasional interruptions in data service provider operations;
   3) as needed, with interruptions in data service provider operations a secondary concern.

b. The data service provider shall perform systems engineering, test engineering, configuration management, COTS procurement, installation of COTS upgrades, network/communications engineering and other engineering functions as required for performance of its mission.
   Levels of Service:

1) no or very infrequent interruptions of data service provider operations;

2) occasional interruptions in data service provider operations;

3) as needed, with interruptions in data service provider operations a secondary concern.

2.11 Technical Coordination Requirements / Levels of Service

a. The data service provider shall provide staff required for participation in SEEDS processes, including ESE data services architecture refinement and evolution, and information technology planning.

b. The data service provider shall provide staff required for participation in SEEDS processes to coordinate data stewardship standards and practices and development and maintenance of standards for content of life cycle data management plans.

c. The data service provider shall provide staff required for participation in SEEDS processes to coordinate best practices among ESE data service providers, including quality assurance standards and practices for all phases of data services provider functions.

d. The data service provider shall provide staff required for participation in SEEDS processes, and cooperating with other ESE data service providers in representing ESE / SEEDS in broader community processes, for developing and maintaining common standards and interface definitions, including those that enable interoperability within the ESE / SEEDS environment and with other systems and networks as needed to support the ESE program.

e. The data services provider shall participate in SEEDS level and/or bilateral processes to coordinate production and delivery of products between ESE data service providers.

f. The data services provider shall participate in SEEDS processes for coordinating user support guidelines and practices among ESE data services providers.

g. The data services provider shall provide staff required for SEEDS coordination of security standards and practices to meet NASA or other established security requirements.

h. The data service provider shall provide staff to coordinate standards for common metrics.

i. The data service provider shall provide funding for travel to support technical coordination activities.

2.12 Implementation Requirements / Levels of Service

a. The data service provider shall design and a data and information system capable of meeting its mission requirements. The design shall address hardware configuration and interfaces and allocation of function to platform. The design shall address software configuration, including COTS, software re-use, and new custom software to be developed, including science software embodying product generation algorithms and/or software facilitating integration of science software provided by outside source(s).

b. The data service provider shall develop a staffing plan that addresses staff required to implement and operate the data service provider over its planned lifetime. The staffing plan shall include a breakdown of positions and skill levels assigned to functions.

c. The data service provider shall develop a facility plan, including planning for space, utilities, furnishings, etc., required to support its staff, data and information system, data storage, etc., and the environmental conditioning to be provided.

d. The data service provider shall accomplish the implementation of its data and information system, including purchase and installation of hardware, purchase or licensing and installation and configuration of COTS software, modification, installation and configuration of re-use software, development of new custom software, and integration of all components into a tested system capable of meeting the data service provider's mission requirements.

e. The data service provider shall perform ongoing applications software development.
   Levels of Service:

   1) Below plus implementation of applications software to perform a 'data mining' or data integration operation to meet a user need.

   2) Below plus implementation of product generation software embodying science algorithms, e.g. to produce a product to meet a particular user need;

   3) Implementation of software tools for use by users to unpack, subset, or otherwise manipulate products provided by the data service provider;

f. The data service provider shall provide the staff needed to accomplish all needed in-house development and test activities.

2.13 Management Requirements / Levels of Service

a. The data service provider shall provide management and administrative staff to perform supervisory, financial administration, and other administrative functions.

b. The data service provider shall provide staff required for participation in SEEDS management processes, strategic planning, coordination with other data centers and activities beyond ESE/SEEDS.

c. The data service provider shall provide staff with science expertise to coordinate the science activities within the data service provider and its interaction with the ESE and broader science community, including a visiting scientist program (or equivalent), collaboration among ESE data service providers to support science needs, annual Enterprise peer review, and support for its User Advisory Group and any other advisory activities appropriate given its ESE role and user community.

d. The data service provider shall provide staff with system engineering expertise to plan information technology upgrades / technology refreshes, based on assessments of changing mission or user needs and availability of new technology. (Coordination with other ESE data service providers is included in technical coordination).

e. The data service provider shall provide staff with data management expertise to develop data stewardship practices, perform data administration with science advice (via the User Advisory Group and other appropriate bodies), develop and maintain life cycle data management plans including data migrations. (Coordination with other ESE data service providers is included in technical coordination).

2.14 Facility / Infrastructure Requirements / Levels of Service

a. The data service provider shall maintain site, system, and data security according to established NASA or other policies and practices while providing easiest possible access (consistent with required security) to its data and information services for its user community.

b. The data service provider shall provide and maintain a fully furnished and equipped, environmentally controlled, physically secure facility to house its staff, systems, and data and information holdings.

c. The data service provider shall provide a backup facility for its data and information holdings.
   Levels of Service:

   1) an environmentally controlled and physically secure off-site backup archive facility;

   2) an on-site but separate environmentally controlled and physically secure off-site backup facility;

   3) a backup capability within the data service provider's primary data system(s).

d. The data service provider shall perform resource planning, logistics, supplies inventory and acquisition, and facility management.
   Levels of Service:

1) no or very infrequent interruptions of data service provider operations;

2) occasional interruptions in data service provider operations;

3) as needed, with interruptions in data service provider operations a secondary concern.

e. The data service provider shall provide network connections and services as needed to support its operations.

3 User-Oriented View of Levels of Service

This section presents a user-oriented view of the requirements / levels of service given in Section 2. The intent is to provide users with a description of data service provider services and levels of service that they would actually see or interact with. The goal is to provide users with a description of services in terms of their experience, to facilitate their critique and review.

Users in this context include other data service providers (such as applications centers who need ESE products as input, or flight projects who provide data to a data service provider for archive and distribution and so interact with the provider's ingest service) as well as 'end users' such as research scientists or applications specialists who interact with the provider's search and order and access and distribution services.

Because the focus of this section is on the user's interaction with the data service provider, services that the user does not directly interact with, such as sustaining engineering or facility support, are not included, even though successful performance of these services is essential to the success of the services the user actually sees. Services that the user actually sees or interacts with are ingest, processing, documentation, search and order, access and distribution, and user support.

The user-oriented service descriptions are organized according to the reference model's functional areas. Each service is presented in a table containing from one to five levels of service, ranging from 'lowest' through 'low', 'medium', 'high', to 'highest'. In some cases the increasing levels of service will be additive or cumulative, i.e. the next highest service will say "Add: [service]" meaning that at this level of service the [service] is to be added to the service available at the level immediately below it.

The table also includes a reference to the 'engineering form' of the service description in Section 2 above. As noted above, not all of the 'engineering form' requirements / levels of service will map to user-oriented service descriptions in this section. There can also be cases where more than one user-oriented service description will be mapped to a requirement / levels of service in Section 2. These are cases where facets of a service that may be of significance from the user point of view are not separable or resolvable as separate requirements / levels of service in the reference model for cost estimation.

The levels of service are presented in the context of the general data service provider reference model. An evaluation of which services at what levels are applicable to an actual real-world data service provider would depend on an analysis of its ESE role and mission, the needs of its user community, the data it supports, etc. A general indication of how the services and levels of service will vary is provided in White Paper 6, "ESE Logical Data Service Provider Types".

3.1 Ingest Service

This section presents a users' view of ingest services provided by the data service provider. A user of a DSP's ingest service might be another provider that generates products and delivers them to the DSP to be ingested, archived, and distributed.

1. Product Ingest Service

Level	Service - What does the user see?	Reference: 2.1 a
Lowest	The DSP will occasionally accept products on an ad-hoc basis.	
Low	The DSP will accept products on an ad-hoc basis.	
Medium	The DSP will accept products on an ad-hoc basis and will return verification of receipt and successful ingest (i.e., including verification of data quality and integrity).	
High	The DSP will accept products on a routine basis and will return verification of receipt and successful ingest.	
Highest	The DSP will accept products on an operational, time-critical basis and will return immediate verification of receipt and successful ingest.	

In practice, the requirement for a particular level of service for ingest will vary from data stream to data stream; a real-world data service provider will ingest multiple data streams or flows of products at different levels of service. A distinction between data service providers will be the highest level of service they will provide for any data stream.

3.2 Processing Services

This section presents a users' view of the processing, or product generation, services provided by a data service provider. Users of a DSP A's processing service might include:

1) another DSP whose own product generation might be dependent upon products being generated on a timely basis by DSP A;

2) a flight project team who 'subcontracted' routine generation of products from its mission data to DSP A;

3) a research project team who turned to the DSP to generate research products from DSP-held data on an ad-hoc basis;

4) a researcher involved with a field experiment who depends on DSP products as aids in conducting the field campaign.

In the second and third cases, the user might wish to provide the science algorithms according to which a product is generated, perhaps retaining responsibility for science quality assurance and for judging the readiness of a product for general use.

In the discussion of processing services, two types of products are referred to, 'standard' products and 'ad hoc' products. Standard products are products that are produced over a period of time using a validated processing algorithm that gains both stability and peer acceptance, such that a science or applications user of the product can rely upon it and use it with confidence. Other products might be produced for a specific study, or might be research products of uncertain quality whose limitations a user would need to take into account.

A data service provider might offer at least three distinct modes of production to its users:

1) operational generation of new products, whether scheduled (perhaps to keep up with data inflow), or on demand, characterized by high reliability and robustness, perhaps with terms of service formally agreed to a level of service agreement or equivalent;

2) non-operational, where new products are generated without a fixed schedule or guarantee of responsiveness, perhaps using DSP resources on an as available basis;

3) reprocessing, generating new version(s) of previously generated products, either on request or according to a schedule negotiated with the user (where the user might be a flight project or science team responsible for the science algorithm software used to generate the products.

1. Operational Product Service

Level	Service - What does the user see?	Reference: 2.2 a
Lowest	The DSP will produce a product within 30 days of receiving its inputs.	
Low		
Medium	The DSP will produce a product within 7 days of receiving its inputs, by a schedule or in response to an on-demand request.	
High		
Highest	The DSP will produce a product within 2 days of receiving its inputs, by a schedule or in response to an on-demand request.	

The operational product service is generally associated with standard products. A real-world data service provider will provide a mix of levels, depending on the requirements for each of its products. A distinction between data service providers will be the highest level of service they will provide for any production stream.

2. Non-Operational Product Service

Level	Service - What does the user see?	Reference: 2.2 b
Lowest	The DSP will produce non-operational products on a time available basis - no goals or targets.	
Low		
Medium	The DSP will produce non-operational products, meeting general goals negotiated with the user.	
High		
Highest	The DSP will produce non-operational products, meeting specific targets negotiated with the user.	

3. Reprocessing Service

Level	Service - What does the user see?	Reference: 2.2 c, d
Lowest	The DSP will reprocess products per user request, on a time available basis - no goals or targets.	
Low		
Medium	The DSP will reprocess products, meeting general goals negotiated with the user.	
High		
Highest	The DSP will reprocess products, according to a reprocessing schedule negotiated with the user.	

The reprocessing service is generally associated with standard products, where further research causes a processing algorithm to be improved to the point where scientists recommend reprocessing to produce a new version of the product using the improved algorithm. Reprocessing of a product might also be driven by reprocessing of one or more of the inputs used to produce it, or of corrections or improvements to ancillary data.

4. Science Software Integration and Test Service

Level	Service - What does the user see?	Reference: 2.2 e
Lowest	The DSP will not accept user provided software.	
Low		
Medium	The DSP will accept science software from a user for standard products to be produced by the DSP, and perform integration test and verify readiness of the software for production.	
High	Add: The DSP will accept science software from a user for research products to be produced by the DSP, and perform integration test and verify readiness of the software for production.	
Highest	Add: The DSP will accept science software from a user for data integration or data mining tasks, and perform integration test and verify readiness of the software for production.	

3.3 Documentation Service

This section presents a users' view of the product documentation services provided by a data service provider. All users of a DSP's products will rely on the DSP's documentation to support their understanding and use of the products. Users such as a flight project or research project that provides products to a DSP for archive and distribution, or that 'subcontracts' product generation to a DSP, will be concerned with the quality of DSP documentation, especially if members of their team will have to use it. The eventual long term archive for the DSP's data and products will also have a vital concern with the quality of DSP documentation.

1. Documentation Service

Level	Service - What does the user see?	Reference: 2.3 a, b, c
Lowest	The DSP provides GCMD DIFs describing its data and products, and other documentation either only as received from data or product sources or informal documentation of products it generates.	
Low	Add: The DSP provides read software and format documentation for its data and products.	
Medium	Add: The DSP provided users' guides for its data and products (product type level, product instance (a.k.a. granule) level, electronic or hard copy, including journal references).	
High	Add: The DSP includes user feedback about data and products in its documentation.	
Highest	Add: The DSP ensures that its data and products are documented to the standard required for long term archiving.	

3.4 Archive Services

This section presents a users' view of the archive service provided by a data service provider. For many DSP users, especially researchers or applications specialists who obtain data or products from the DSP, the archive service will be out of sight, a layer below the access services they interact with directly. They will rely implicitly on the archive service as they do on the DSP's sustaining engineering and other vital but not directly visible services. Other users of the DSP's archive service will include flight projects or research teams who will rely on the DSP to archive their data and products, and while they would accomplish delivery of data and products to the DSP via the DSP's ingest service, they will be concerned with the level of archive service the DSP will provide for their data and products. What would be visible to the user in this sense is the commitment of the DSP to provide a given level of service (perhaps captured in a formal agreement).

1. Archive Quality Monitoring Service

Level	Service - What does the user see?	Reference: 2.4 e, f
Lowest	Quality screening (read after write) on data entering the archive;	
Low	Add: 1% per year random screening to detect and replace degrading media;	
Medium	Add: 5% per year random screening;	
High	Add: Exit screening (tracking of archive read failures, corrected errors, etc., to find and replace degrading media);	
Highest	Add: 10% per year random screening;	

A user who would see reports of the quality monitoring might be a flight project or research group that had entrusted the data service provider with the stewardship of data critical to its work. A 'regular' end user would not see the effects of this service directly. The next service, archive backup service, is a similar case.

2. Archive Backup Service

Level	Service - What does the user see?	Reference: 2.4 h, i
Lowest	Partial on-site backup, with regular sampling to verify integrity of the backup.	
Low	Add: Media migration on ad hoc basis when needed.	
Medium	Add: Partial off-site backup, with regular sampling to verify integrity of the backup.	
High	Add: Full off-site backup, with regular sampling to verify integrity of the backup.	
Highest	Add: Media, archive technology migration planned and budgeted for.	

3.5 Search and Order Services

This section presents a users' view of the search and order service provided by a data service provider.

1. Availability of Search and Order Service

Level	Service - What does the user see?	Reference: 2.5 a
Lowest	Search and Order service is open to a limited team of scientists or applications specialists.	
Low		
Medium	Search and Order service is open to the science and/or an applications community.	
High		
Highest	Search and Order service is public, open to all users.	

2. Search Service

Level	Service - What does the user see?	Reference: 2.5 b
Lowest	Search a list of available DSP products for ones of interest.	
Low	Search for DSP products of a specified type by time and space coverage.	
Medium	Search for any/all DSP products by time and space coverage.	
High	Add: search for any/all DSP products by geophysical parameter, time, and space coverage.	
Highest	Add: search for any/all DSP products pertaining to a given phenomenon (e.g. hurricane, volcanic eruption, El Nino), campaign, research project.	

3. Search Service - Product Descriptions

Level	Service - What does the user see?	Reference: 2.5 c
Lowest	Option to view standard guide and DIF metadata for any product found by a search.	
Low	Add: option in some cases to view a few references to published papers that describe the production or use of the product.	
Medium	Add: option in most cases to view a few references to published papers that describe the production or use of the product.	
High	Add: option in some cases to view detailed algorithm and use explanations for the product.	
Highest	Add: option in most cases to view detailed algorithm and use explanations for the product.	

3. Search and Order Service Mode

Level	Service - What does the user see?	Reference: 2.5 b, d
Lowest	Web accessible user interface for search and order for the DSP.	
Low		
Medium	Web accessible user interface for search and order for this and other ESE DSPs.	
High		
Highest	User system accessible interface for automated search and order.	

3.6 Access and Distribution Services

This section presents a users' view of the access and distribution service provided by a data service provider.

1. Availability of DSP Data and Product Holdings

Level	Service - What does the user see?	Reference: 2.6 a
Lowest	DSP data and products are available to a limited team of scientists or applications specialists.	
Low		
Medium	DSP data and products are available to the science and applications community.	
High		
Highest	DSP data and products are public, open to all users.	

In the case of a real-world data service provider, product availability can vary on a product by product basis. Some products (such as those obtained from international or commercial sources) might have general

distribution restrictions that NASA accepts in order to secure access for approved NASA scientists. Other products might go through a period of limited access while they are validated and corrected or improved before being approved for general access. What may distinguish data service providers from one another is the highest level of availability they are willing or required by their ESE role to support.

2. Access and Distribution Service Mode

Level	Service - What does the user see?	Reference: 2.6 d, e
Lowest	Data available on a provider scheduled basis, no flexibility for user.	
Low	Data available to user on a request basis.	
Medium	Add: data available to user on a subscription (i.e. standing order) basis.	
High	Add: data available to user on an operational basis (DSP will commit to terms of service for scheduled or on demand distribution).	
Highest	Add: user's system can access data from DSP system directly by network.	

3. Data Services

Level	Service - What does the user see?	Reference: 2.6 c
Lowest	Reformatting available for a few selected DSP products.	
Low	Reformatting, subsetting, available for less than half of DSP products.	
Medium	Subsetting, reformatting, packaging available for less than half of DSP products.	
High	Subsetting, reformatting (including GIS support), packaging available for more than half of DSP products.	
Highest	Subsetting, reformatting (including GIS support), resampling, reprojection, packaging available for most DSP products.	

The specific data services available for a particular product will depend on the nature of the product and the needs of its users; the higher levels of service mean a progressively wider variety of data services available for a progressively increasing fraction of the data service provider's holdings.

4. Access and Distribution Service Timeliness

Level	Service - What does the user see?	Reference: 2.6 f, g
Lowest	Availability of a product for network delivery (e.g. staged for FTP pick up or push) within twenty four hours of request (or of production of a product ordered in advance). Shipping of a product on one form of media within one month of request.	
Low	Add: Shipping of a product on media, with user's choice from several media types, within one week of request.	
Medium	Add: Availability of a product for network delivery within ten minutes.	
High	Add: Availability of a product for network delivery within ten seconds (e.g. data or products held on-line). Shipping of a product on media within three days of request.	
Highest	Add: Availability of a product for automated, direct access via network within ten seconds (e.g. data or products held on-line for access by user software).	

In the case of a real-world data service provider, the level of service will almost always vary with the product, depending in each case on factors such as user demand and sheer size. Also, these service levels are defined in terms of a request for a single product. A request for a year's worth of a product, especially a large low level product, e.g. MODIS level 1, would take much longer, and would be negotiated with the user.

3.7 User Support Services

This section presents a users' view of the user support service provided by a data service provider.

1. Availability of User Support

Level	Service - What does the user see?	Reference: 2.7 c, d
Lowest	Casual telephone or email contact.	
Low	Add: Help Desk staffed during local work day (5 days x 8 hours per day).	
Medium	Add: On-line help, FAQ, data / product and service descriptions.	
High	Add: Help Desk staffed during work day for all U.S. (5 days x 12 hours/day).	
Highest	Add: Help Desk staffed during work day worldwide (5 or 7 days x 24 hours per day).	

2. Capability of User Support Staff

Level	Service - What does the user see?	Reference: 2.7 b
Lowest	Basic knowledge of what data and products are available, what network / media / delivery options exist.	
Low	Add some knowledge of format detail for most popular products.	
Medium	Add comprehensive knowledge of format details for most if not all products.	
High	Add technical expertise in data structures, use of tools for format conversion, subsetting, analysis, etc..	
Highest	Add science expertise in data / product quality and research uses for data and products.	

3. Outreach to Potential New Users

Level	Service - What does the user see?	Reference: 2.7 e
Lowest	None.	
Low	DSP outreach material (pamphlets, brochures, posters, etc.)	
Medium	DSP's booth or booth participation at conferences, at least four times/year.	
High	Add: DSP staff providing or contributing to workshops and/or user training sessions at conferences.	
Highest	Add: DSP staff providing user training sessions at universities, schools, etc.	

3.8 Applications Software Service

This section presents a users' view of the application software service (a subset of implementation as defined above) provided by a data service provider. Users, except possibly for an advisory panel, will not see or interact with the initial implementation of the data service provider's system capabilities or their technology refresh or expansion. Users will see applications software developed by the data service provider to meet users' needs.

1. Applications Software Service

Level	Service - What does the user see?	Reference: 2.12 e
Lowest	None.	
Low	DSP developed software tools to read DSP products.	
Medium	DSP developed software tools for use by users to unpack, subset, or otherwise manipulate DSP products.	
High	Add: DSP developed product generation software embodying science algorithms, e.g. to produce a product to meet a particular user need.	
Highest	Add: DSP developed applications software to perform a 'data mining' or data integration operation to meet a user need.	

Appendix A - Draft Program Level Requirements

This section contains the set of program level requirements drafted by the SEEDS Formulation Team in September, 2001, as "NewDISS Level 0 Requirements", with the new term "SEEDS" replacing "NewDISS". The cost model requirements template that follows fits within the general framework of the program level requirements in this section.

A.1 General Requirements

a. Data service providers will fully participate (TBD) in SEEDS community-based management processes including standards and interface determination, reuse/architecture refinement, metrics collection, and Enterprise peer review.

b. All data service providers will comply with SEEDS Level of Service requirements for core functions and data products (TBD) and will adhere to SEEDS required core interfaces and standards (TBD). Deviation from core standards must be requested and approved via the SEEDS waiver process (TBD).

c. Data service providers will provide metrics (TBD) on data production and utilization to the SEEDS Office on a routine (TBD) basis.

d. Data service providers and projects will participate in an annual (TBD) broad-based peer review of ESE data management activities.

e. ESE Mission Projects will produce a Life Cycle Data Management (LCDM) Plan. Changes to the LCDM plan will be approved by the SEEDS Office (TBD).

f. To the extent possible and where cost effective, data service providers will reuse software and system components developed by previously NASA funded activities. Projects will enable possible reuse of their software available by following the system design guidelines provide by the SEEDS reference architecture (TBD).

A.2 General Science Requirements

a. Data service providers will provide support to and receive technical direction from an appropriate NASA ESE science parameter team.

b. Principal Investigators will propose a suite of standard science products subject to peer review approval of an Algorithm Theoretical Basis Document.

c. Each data service provider will have a Science Advisory Group that will review progress and plans on a routine basis.

A.3 Production, Archive, and Distribution Requirements

a. All raw data will be acquired will be calibrated and geolocated to a reference sphere. Calibrated and georeferenced data will made available to all users.

b. Data at the "raw" sensor level (NASA Level 0 plus appended calibration and geolocation information) must be archived permanently.

c. All standard science data (Level 1b, Level 2, and Level 3) produced will be made available to any user who requests it without discrimination.

d. All standard data products available to a science team member will be made available to general science users.

e. All standard data produced will be archived until the end of the science mission or until transfer to an approved permanent archive.

f. Data service providers will receive orders for data products from the general public and will fulfill those orders with an average delivery time (elapsed time between when the order was completed and product was shipped) of less than five working days.

A.4 Standards and Interface Compliance

a. Metadata for all standard products will be produced in accordance with the SEEDS core metadata standard.

b. Metadata for all archived standard data products must be searchable by spatial and temporal extent, and must be locatable by the general user via the world wide web.

c. Standard data products made available to the LTA, to another SEEDS data service provider and to users will be available in one of the SEEDS core formats.

d. All standard data products will be cataloged in the Global Change Master Directory (GCMD). Data service providers will provide Directory Interchange Format (DIF) documents on all standard data products to the GCMD prior to release of the data products.

References and Acronym List

The References Section and the Acronym List for all of these Working Papers is in the document "References and Acronyms for the Levels of Service / Cost Estimation Working Papers ".